

Measurement of Single-Top Quark Production at the Tevatron

Thomas R. Junk
Fermi National Accelerator Laboratory

on behalf of the CDF and D0 Collaborations

ICHEP 2008, Philadelphia, PA July 30, 2008

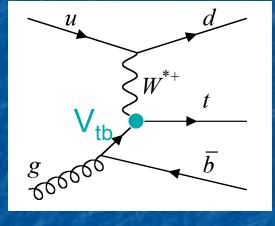
Motivation: Production Rate is Proportional to $|V_{tb}|^2$

$$\sigma_t = (1.98 \pm 0.25) |V_{tb}|^2 \text{ pb}$$

$$\sigma_s = (0.88 \pm 0.11) |V_{tb}|^2 \text{ pb}$$

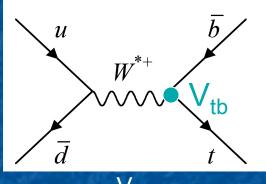
B.W. Harris et al., Phys. Rev. D66, 054024 (2002).
Z. Sullivan, Phys. Rev. D70, 114012 (2004).
Compatible Results;
Campbell/Ellis/Tramontano, Phys. Rev. D70, 094012 (2004).
N. Kidonakis, Phys. Rev. D74, 114012 (2006).

"tqb"



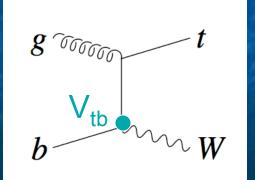
t-channel production

"tb"



s-channel production

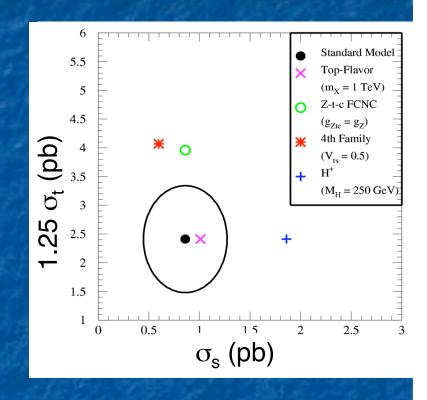
$$g + b \rightarrow W^{-} + t$$
$$g + u \rightarrow t + \overline{b} + d$$



Associated production (small, ~ 0.3 pb)
-- neglected here.
Appears in W+3-jet events.

Physics Motivations for Seeking Single Top

- Interesting signal -- s and t-channel rates are differently sensitive to new interactions
- Single top quarks are ~100% polarized in the SM
 - Can test this with angular distributions of decay products
- A check of the b PDF of the proton
- Exotic couplings -- see also L. Li's talk



T. Tait and C. P. Yuan, Phys.Rev.**D63**:014018 (2001)

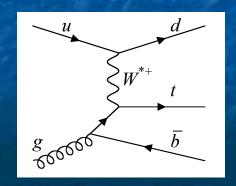
Technical Motivations for Seeking Single Top

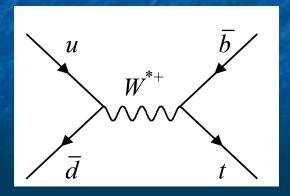
- Single top is a background to WH→Ivbb
 Measurements are preferable to MC predictions
- Its backgrounds are backgrounds to WH→Ivbb (W+jets, ttbar, QCD, dibosons)
- We need to observe it on our way to observing a light Higgs boson
 - It has a larger cross section than WH→Ivbb (order of magnitude)
 - The kinematic signature is more distinct than WH→Ivbb
 - It's a great testing ground for making a discovery using advanced signal/background separation techniques

General Strategy for Measuring Single Top Production



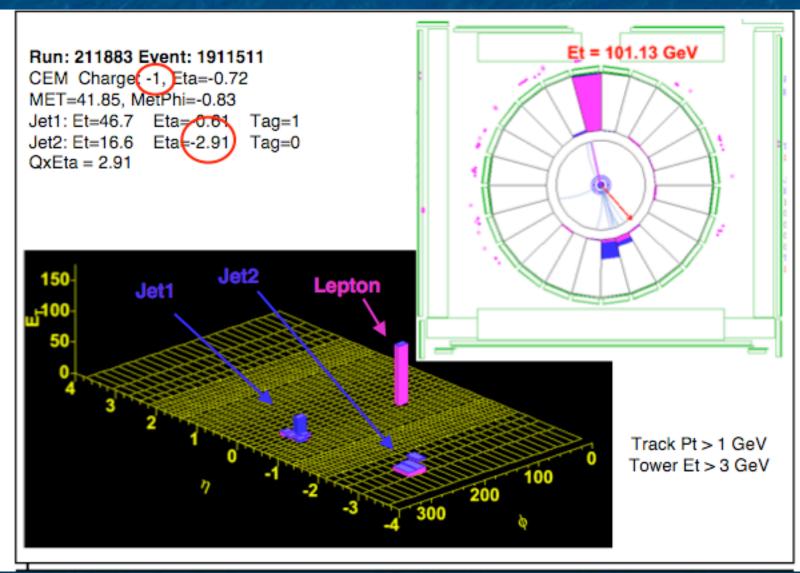
- Model signal with a Monte Carlo which reproduces NLO expectations
- Model Backgrounds with Pythia, ALPGEN, and data control samples
- Select events with basic features expected of single top events
 - A high-P_T lepton (20 GeV or more (CDF), 15 GeV or more (D0))
 - Large Missing-E_⊤ (25 GeV or more (CDF), 15 GEV or more (D0))
 - Two or more jets (typically 20 GeV or more) -- Use high-η jets!
 - One or more b-tags
 - Veto Z→leptons, cosmics, conversions
- Train multivariate discriminants to separate signals from backgrounds
 - Train separately for each number of jets and tags
- Fit Templates to data and extract cross sections and significances
- Check! Check! and Re-Check!

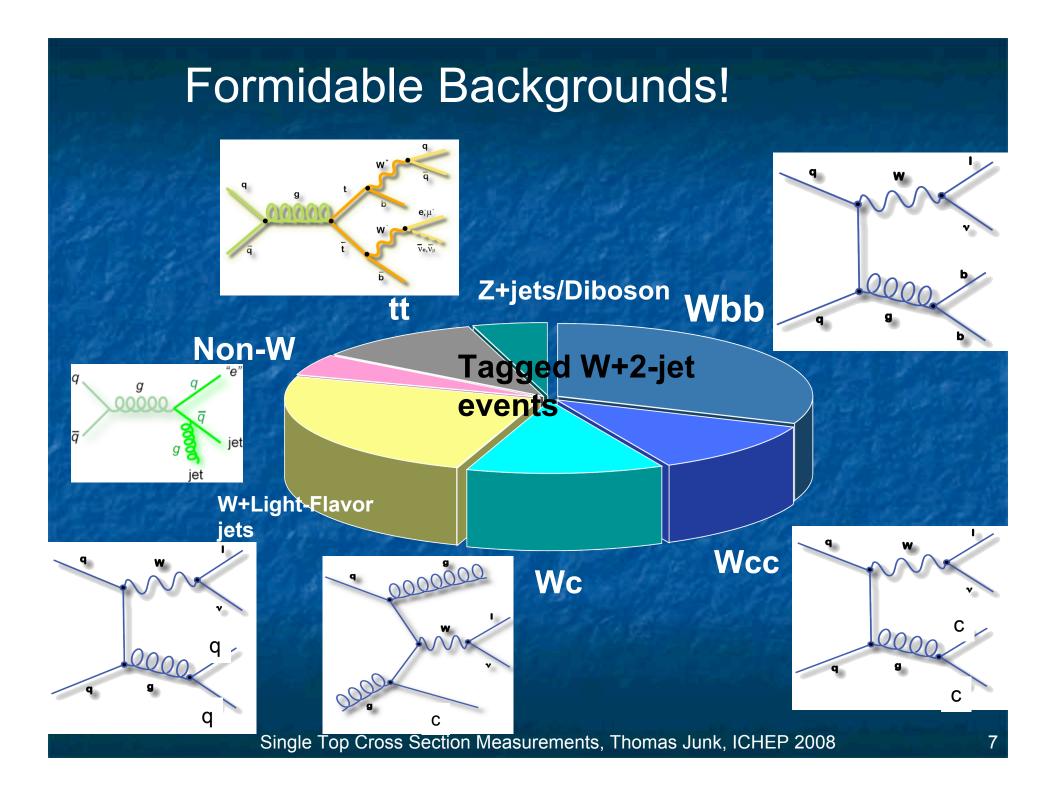






An Event that Looks Like Single Top



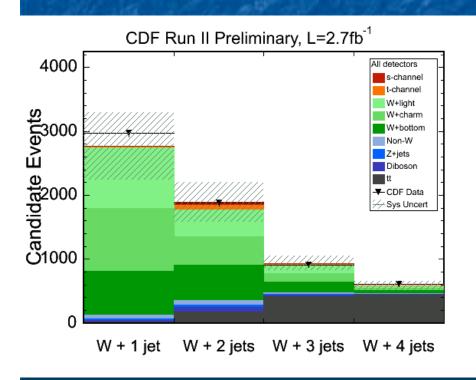


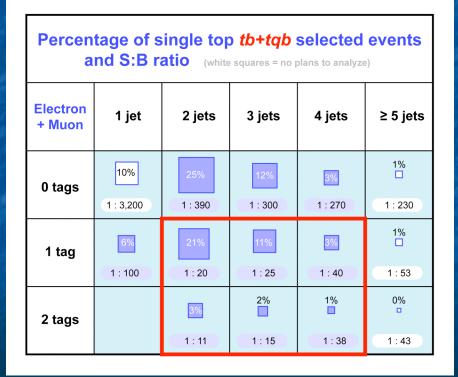


Backgrounds and signal in 2.7 fb⁻¹ by number of jets



V.M. Abazov et al., Phys. Rev. Lett **98**:181802 (2007) V.M. Abazov et al., Phys. Rev. D78:012005 (2008) NEW!







Predicted Rates in 0.9 fb⁻¹

- W+jets simulated with ALPGEN+Pythia, scaled to data. Wbb,Wcc scaled by α =1.5±0.45 after measurement in control samples
- Non-W fraction estimated using loose lepton and low Missing-E_T control samples -- data based models for kinematics
- · top pairs modeled with MC

	Event Yields in 0.9 fb ⁻¹ Data Electron+muon, 1tag+2tags combined		
Source	2 jets	3 jets	4 jets
tb	16 ± 3	8 ± 2	2 ± 1
tqb	20 ± 4	12 ± 3	4 ± 1
tt̄ →	39 ± 9	32 ± 7	11 ± 3
<i>tt̄ → l</i> +jets	20 ± 5	103 ± 25	143 ± 33
W+bb̄	261 ± 55	120 ± 24	35 ± 7
W+cc̄	151 ± 31	85 ± 17	23 ± 5
W+jj	119 ± 25	43 ± 9	12 ± 2
Multijets	95 ± 19	77 ± 15	29 ± 6
Total background	686 ± 41	460 ± 39	253 ± 38
Data	697	455	246

Signal

Backgrounds

SM Rate Estimates in 2.7 fb⁻¹

CDF Run II Preliminary					
Predicted 2-jet event yield with 2.7 fb-1					

s-channel	49.3 ± 7.0	
t-channel	74.3 ± 10.9	
Single top	123.6 ± 17.9	
tt	173.5 ± 24.8	
Diboson	74.5 ± 7.5	
Z + jets	31.1 ± 4.6	
W + bottom	549.1 ± 165.5	
W + charm	453.5 ± 139.9	
W + light	410.7 ± 51.0	
Non-W	75.6 ± 30.2	
Total background	1768.0 ± 311.9	
Total prediction	1891.6 ± 312.4	
Observed	1874	

CDF Run II Preliminary Predicted 3-jet event yield with 2.7 fb⁻¹

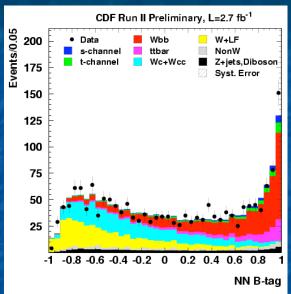
	THE RESERVE OF THE PARTY OF THE	
s-channel	16.3 ± 2.3	
<i>t</i> -channel	22.3 ± 3.2	
Single top	38.6 ± 5.5	
tt	410.5 ± 58.4	
Diboson	25.6 ± 2.6	
Z + jets	13.4 ± 2.0	
W + bottom	169.8 ± 51.3	
W + charm	126.7 ± 39.0	
W + light	125.5 ± 15.8	
Non-W	27.4 ± 11.0	
Total background	898.9 ± 108.2	
Total prediction	937.5 ± 108.3	
Observed	902	



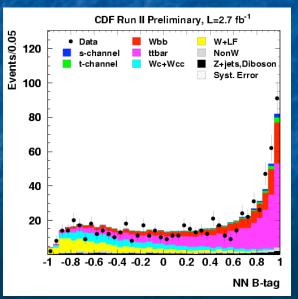
B-Tag Neural Network

~Half of all tagged jets in the W+2jet sample are not b-jets!

- b-tag NN Operates on vertex-tagged jets to separate out charm and light-flavored jets
- · 25 input variables including
 - # tracks with d₀ significance > 3σ
 - Individual track d₀ significance
 - vertex mass
 - lepton p_T (if any)
 - L_{xy} significance and raw L_{xy}

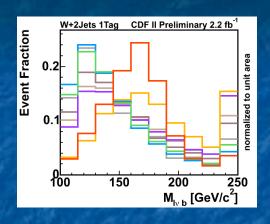


W+2-jet events

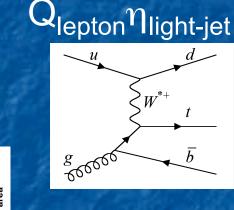


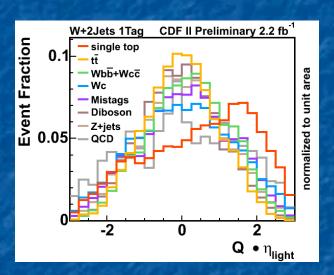
W+3-jet events

Other Important Discriminating Variables



Invariant mass of top decay products. Let's bump-hunt!





W+2Jets 1Tag CDF II Preliminary 2.2 fb⁻¹

- single top
- tt
- Wbb+Wcc
- Wc
- Mistags
- Diboson
- Z+jets
- QCD

0.1

0.1

0.1

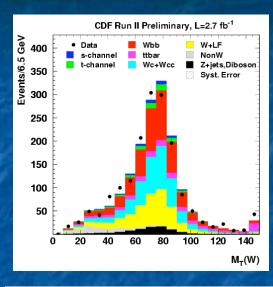
150
200

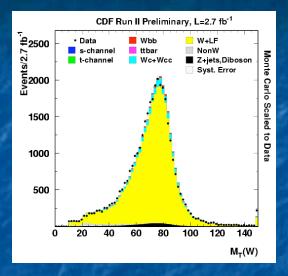
M_{i1i2} [GeV/c²]

m_{ii} -- Distinguishes against W+jets

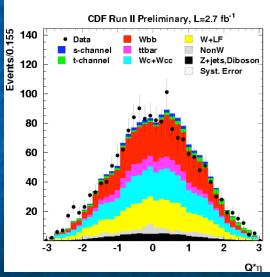
Checking Control Distributions

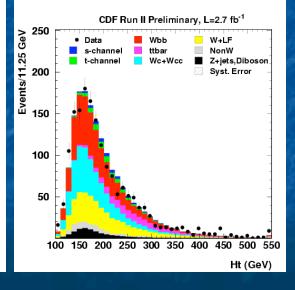








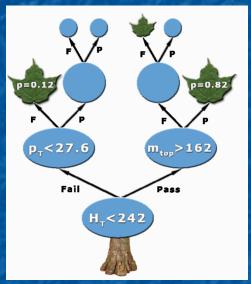


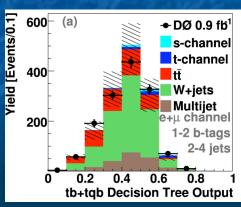


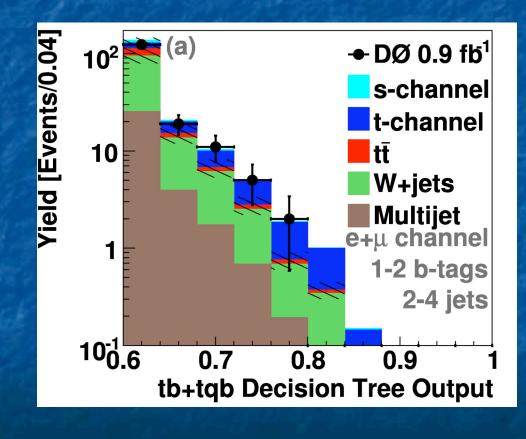
 $Q_{lepton}\eta_{light-jet}$ and H_T = sum of all E_T in event



Multivariate Techniques 1: Boosted Decision Trees



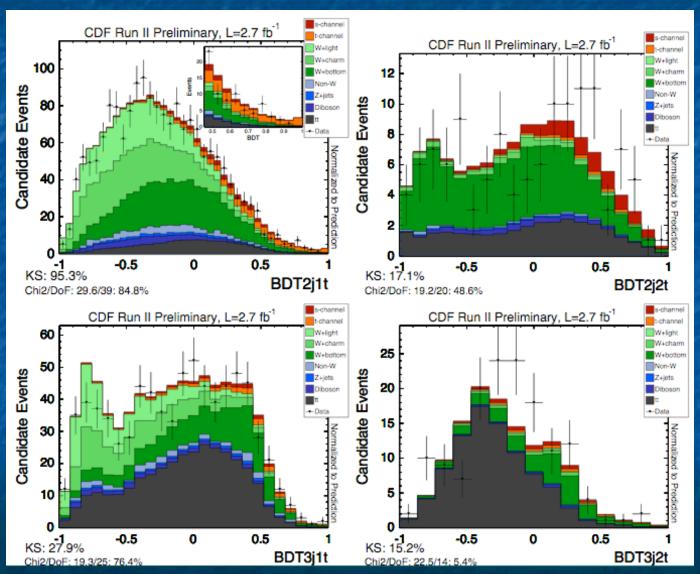




best-fit signal shown: 4.9 pb

Boosted Decision Tree Discriminants







Technique 2: Matrix Elements

Goal -- obtain relative probabilities of each event to come from each process. Same phase space, different matrix elements

Phase space factor: Integrate over unknown or poorly measured quantities

Parton distribution functions

$$P(p_l^{\mu}, p_{j1}^{\mu}, p_{j2}^{\mu}) = \frac{1}{\sigma} \int d\rho_{j1} d\rho_{j2} dp_{\nu}^{z} \sum_{comb} \phi_{4} \left| M(p_i^{\mu}) \right|^{2} \frac{f(q_1) f(q_2)}{|q_1| |q_2|} W_{jet}(E_{jet}, E_{part})$$

Inputs:
lepton and jet 4-vectors no other information
needed!

Matrix element:
Different for each process.
Leading order, obtained from
MadGraph

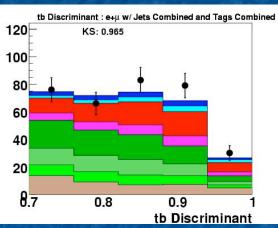
Transfer functions:
Account for
detector effects in
measurement of jet
energy

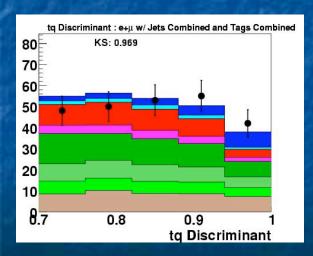
Some processes have no matrix elements, like fakes -- Use a tempate fit like other analyses.

Matrix-Element Discriminants

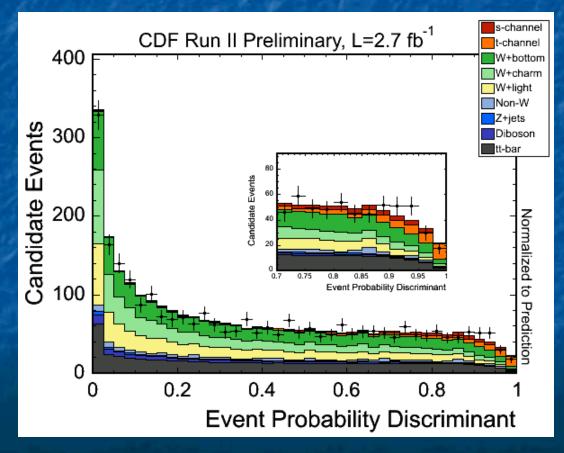








CDF's is optimized together



Technique 3: Neural Networks

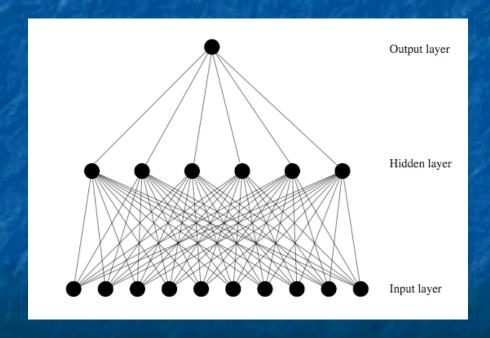
D0: Bayesian Neural Networks

Radford Neal, "Bayesian Learning for Neural Networks", Lecture Notes in Statistics No. 118, Springer-Verlag

CDF: NeuroBayes Program

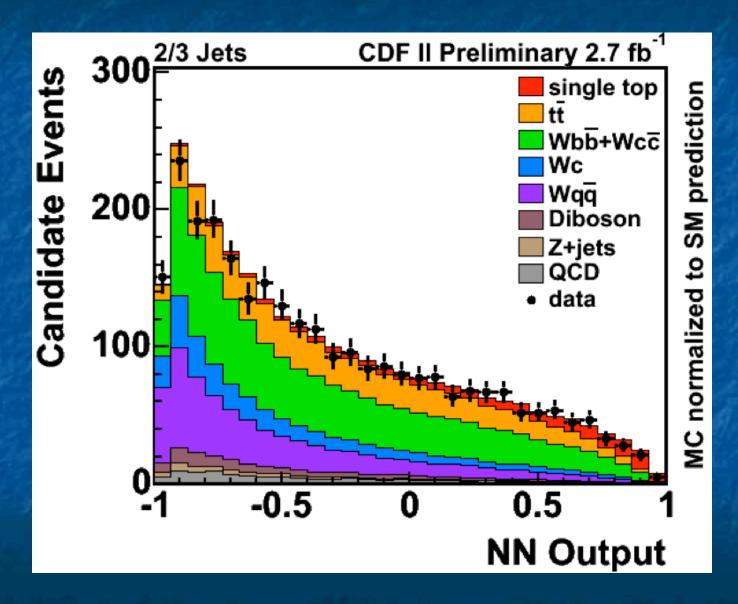
http://www.phi-t.de

- Differ in preprocessing and training.
- Avoids overtraining, even with many input variables



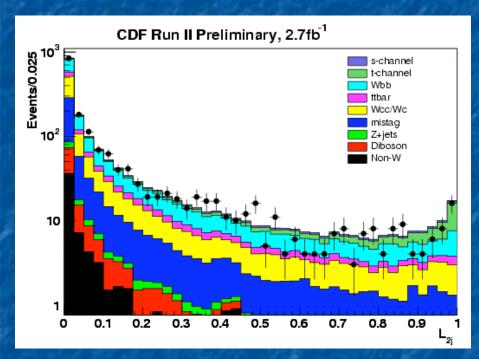
Neural Network Discriminant Distribution





Technique 4: LEP-Style Projective Likelihood Discriminants





CDF Run II Preliminary, 2.7fb⁻¹

s-channel techannel Wbb titbar Wec/We mistag Z+jets Diboson Non-W

2-jet events

3-jet events

Systematic Uncertainty Summary

Source of Uncertainty	Rate	Shape
Jet Energy Scale	016%	✓
Initial State Radiation	011%	✓
Final State Radiation	015%	✓
Parton Distribution Functions	23%	✓
Monte Carlo Generator	15%	
Event Detection Efficiency	09%	
Luminosity	6%	
Neural Net B-tagger		✓
Mistag Model		✓
Q ² scale in ALPGEN MC		✓
Input variable mismodeling		✓
Wbb+Wcc normalization	30%	
Wc normalization	30%	
Mistag normalization	1729%	
ttbar normalization & m _{top}	23%	✓
Non-W Normaliztion	40%	
Non-W Flavor Model		✓

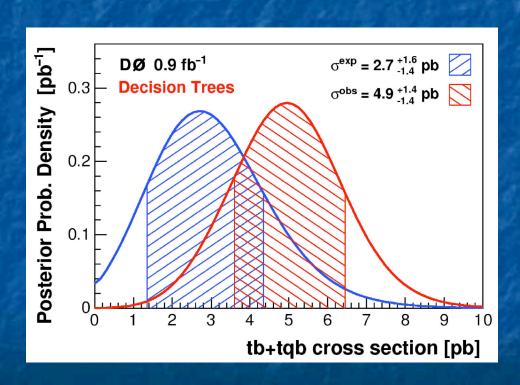


Also:
MC Stats
in each
bin
independently

ISR & FSR treated as correlated

Cross Section Measurements

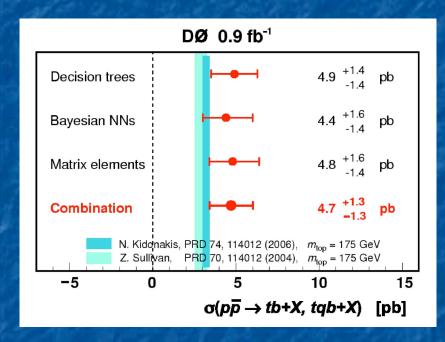
- Both CDF and D0 use Bayesian posteriors integrated over uncertain nuisance parameters
- Flat prior taken in $\sigma_s + \sigma_t$



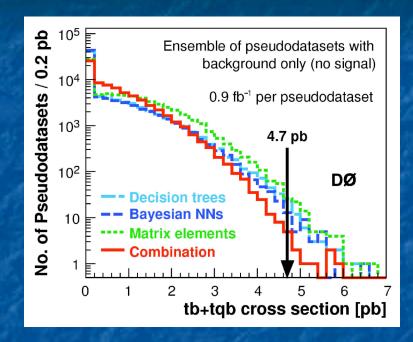
Example -D0 BDT Fit
Excess seen
in data over SM
signal+background
prediciton.



Cross Section Combination and p-Value Calculation -- Evidence for Single Top!

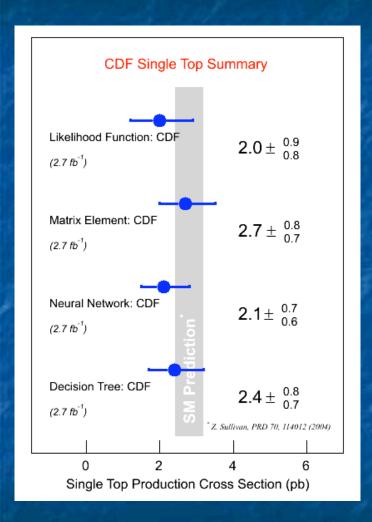


Cross Section measurements combined with BLUE (Best Linear Unbiased Estimate)



Combined cross section measurement used as a test statistic. How many background-only pseudoexperiments would fake a signal this large? A: 1.4×10^{-4} of them. $\rightarrow 3.6\sigma$ evidence. Expected in SM: 2.3σ

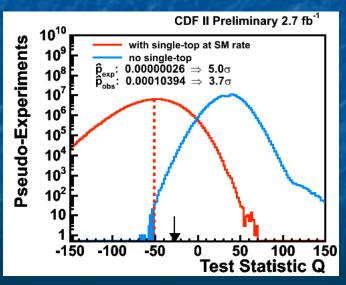
CDF's Cross Sections and P-values



$$Q = \frac{P(\text{data} \mid s + b, \hat{\theta})}{P(\text{data} \mid b, \hat{\theta})}$$

θ=nuisance parameters

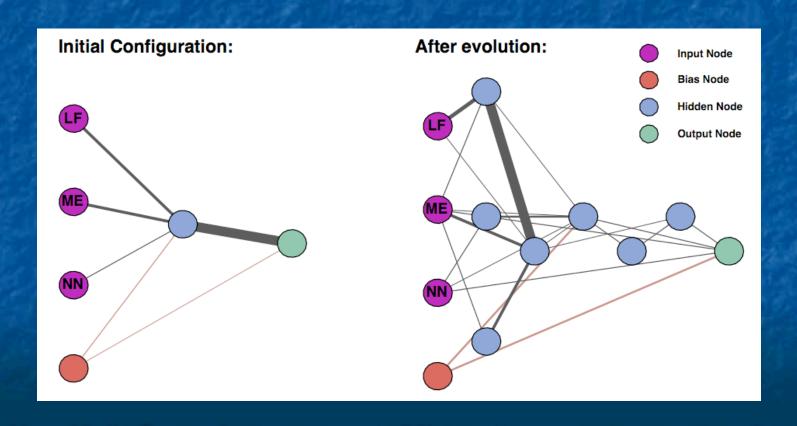
Neyman-Pearson Lemma: Q is the uniformly most powerful test; slightly more powerful than the cross section fit.



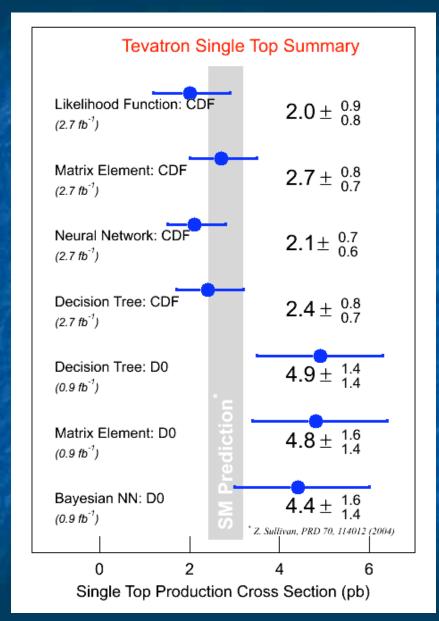
Best Sensitivity: NN analysis 3.7σ observed 5.0σ expected

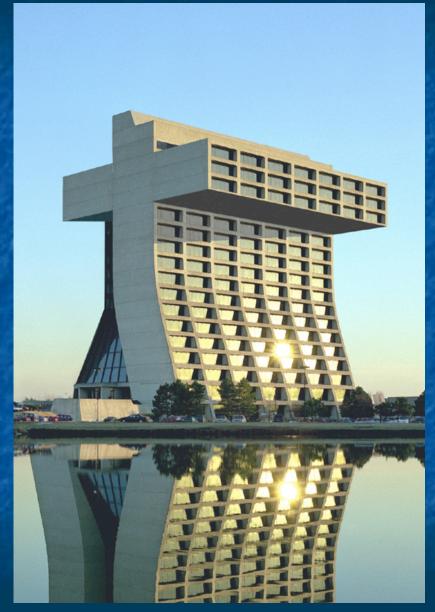
CDF Combination: Super-Discriminant

- Evolutionary neural networks trained to give the best expected p-value, not an arbitrary error function
- Gained CDF 10% in significance for Moriond Combination



Summary of Measurements of σ_s + σ_t

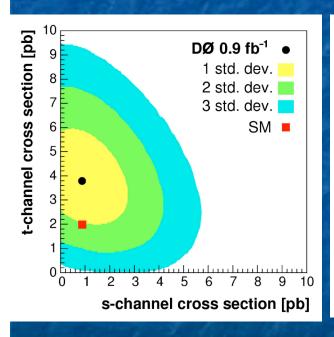


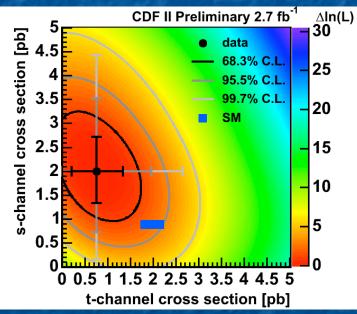


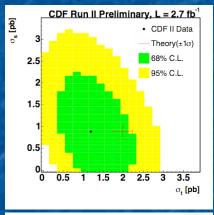
Two-Dimensional Interpretation: Measuring σ_s and σ_t Separately

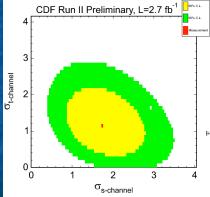
Template fit (marginalization) is done in 2D. Flat priors

taken in the (σ_s, σ_t) plane.









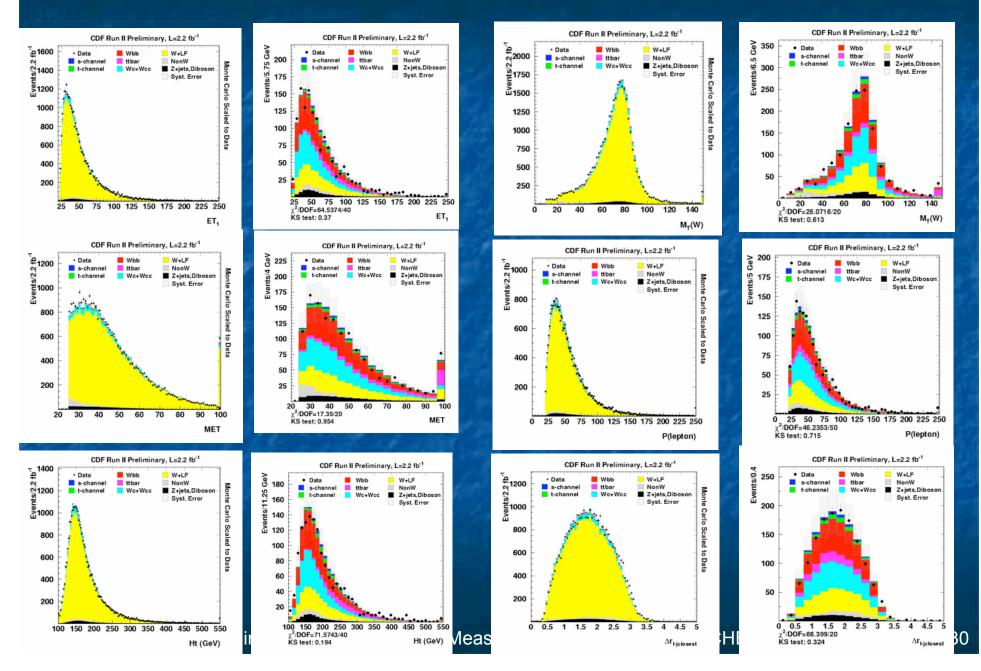
CDF: Dedicated NN's/LF's trained for the purpose of 2D fits.

Summary

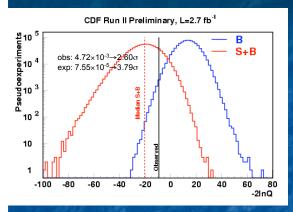
- It's an exciting time to be doing single top!
- D0: Evidence for single top production in 0.9 fb⁻¹ of data. $\sigma_s + \sigma_t = 4.7 \pm 1.3$ pb Significance = 3.6 σ Expected: 2.3 σ
- CDF: Evidence for single top production in 2.7 fb⁻¹ of data. $\sigma_s + \sigma_t = 2.0$ -- 2.7 pb Errors are \pm 0.7 pb per analysis. Expected sensitivity 5.0 σ (Best single analysis) Observed significance: 3.7 σ
- Two-Dimensional contours measured

Backup Slides

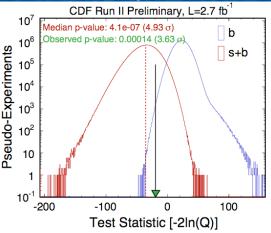
Just Some of the Distributions Checked



P-Value Calculations

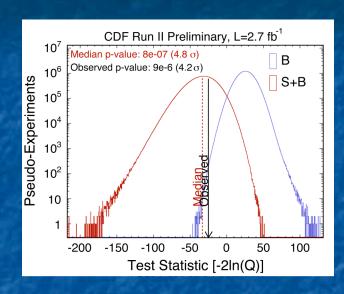


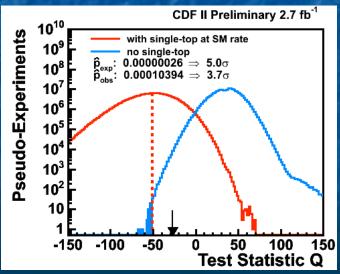
LF: obs: 2.6σ exp: 3.8σ



BDT: obs: 3.6σ

exp: 4.9σ







ME: obs: 4.2σ exp: 4.8σ

NN: obs: 3.7σ exp: 5.0σ



Measurement of Single Top Production in 2.7 fb⁻¹ of Tevatron Data

Update of 2.2 fb⁻¹ measurement presented at Moriond 2008

- Similar Event Selection As D0.
 - Harder Missing E_T cut (25 GeV instead of 15 GeV)
 - Harder jet E_T cuts (all at 25 GeV; D0: 25, 20, 15 GeV)
- W+Jets rates scaled to data in untagged sample
- W+HF Fraction in Alpgen calibrated in W+1-jet sample HF k-factor (D0's α) = 1.4 ± 0.4 for Wbb and Wcc
- b-tagged W+LF calibrated with a tag-rate matrix times pretag W+jets, minus HF, ttbar and non-W
- Non-W Fraction fit in each jet, tag category using Missing E_T distributions

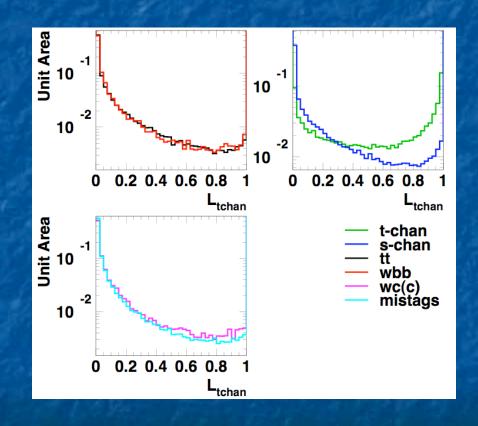
LEP-Style Projective Likelihood Function Discriminant



2-jet bin input variables:

- H_T
- $\cos \theta_{
 m lepton, other jet}$
- Q×η
- M_{ii}
- log(ME_{t-chan}) from MADGRAPH (Stelzer and Willenbrock)
- ANN b-tag output
- χ²(t-channel) (replaces m_{lvb})

Does not take advantage of correlations



Separate network trained for 3-jet events

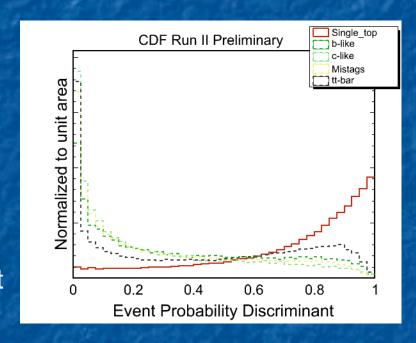
Matrix Element Discriminant



For the same point in phase space, data composition given by ratios of matrix elements squared.

Event probability constructed by integrating matrix element squared times transfer function over unknown parton/neutrino momenta. Use Wbb, Wgg, ttbar, t-channel signal, s-channel signal matrix elements, plus b-tag discriminant

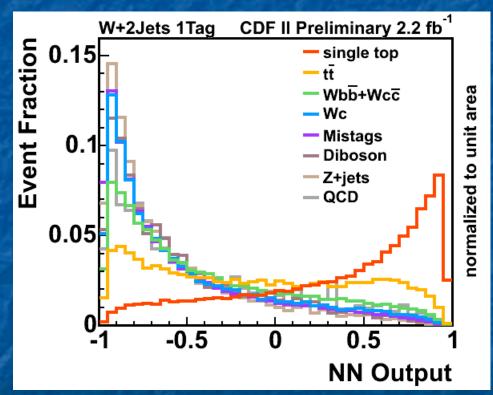
Some components (e.g. non-W) do not have matrix elements - template fits with MC and data-driven models



Neural Network Discriminant



- Networks trained using NeuroBayes
- Four networks trained (2,3 jets)x(1,2 tags)
- CDF's most sensitive analysis



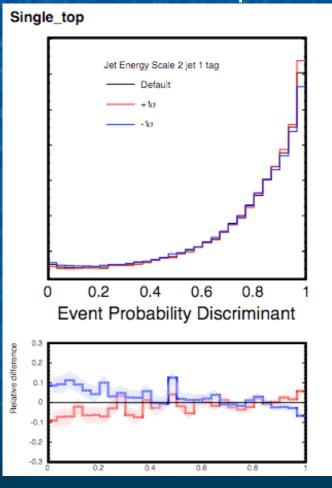
Example templates in 2-jet, 1-tag category

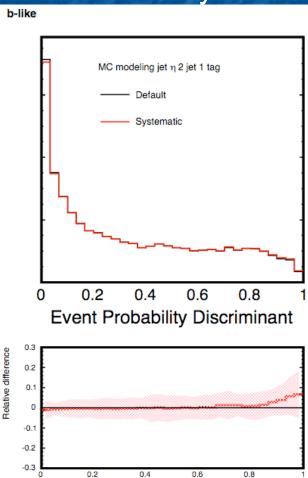
Systematic Shape Uncertainties

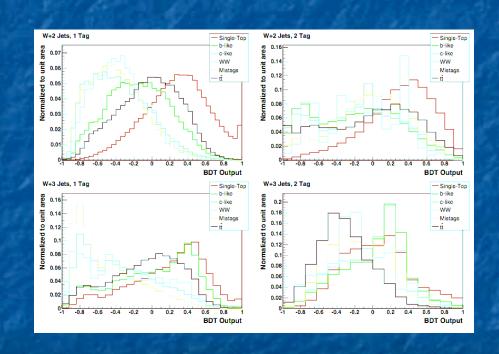
A total of 370 shape uncertainties evaluated!

Most are quite small

Each template, each source of shape error, each channel (#tags, #jets, extra muons) Two examples from the ME analysis







Status of the CKM Matrix

From the PDG review 2006 (Ceccucci, Ligeti, Sakai)

$$V_{\text{CKM}} \equiv V_L^u V_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

$$V_{\rm CKM} =$$

$$\begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272^{+0.0010}_{-0.0010} & (3.96^{+0.09}_{-0.09}) \times 10^{-3} \\ 0.2271^{+0.0010}_{-0.0010} & 0.97296^{+0.00024}_{-0.00024} & (42.21^{+0.10}_{-0.80}) \times 10^{-3} \\ (8.14^{+0.32}_{-0.64}) \times 10^{-3} & (41.61^{+0.12}_{-0.78}) \times 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix}$$

Rigid SM prediction → Measurement of |V_{tb}| is a solid test of assumptions

- Magnitudes only
- 3x3 Unitarity enforced

What Can Affect |V_{tb}|?

• 3x3 Unitarity is not true if the matrix is 4x4 (or more..)

$$V_{CKM} = \begin{vmatrix} V_{ud} & V_{us} & V_{ub} & V_{uX} ? \\ V_{cd} & V_{cs} & V_{cb} & V_{cX} ? \\ V_{td} & V_{ts} & V_{tb} & V_{tX} ? \\ V_{yd} ? & V_{ys} ? & V_{yb} ? & V_{yX} ? \end{vmatrix}$$

Vector t' quark is also a possibility

J. Alwall et. al., "Is |V_{tb}|~1?" Eur. Phys. J. C49 791-801 (2007).

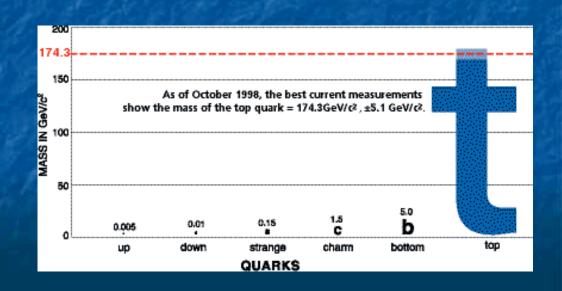
Interesting constraints from precision EW measurements.

Why Measure Single Top Production?

- The top quark is very heavy! Why?
- $m_t = 172.6 \pm 1.4 \text{ GeV/c}^2$ -- Tevatron, March 2008
- Higgs Yukawa coupling ~1

The Top quark may be trying to tell us something!

- Why are other quarks so light?
- Alternative Higgs model t-tbar condensate?
- One of the main ingredients of the radiative corrections to m_H is the top loop
- Is there more out there we cannot yet see?

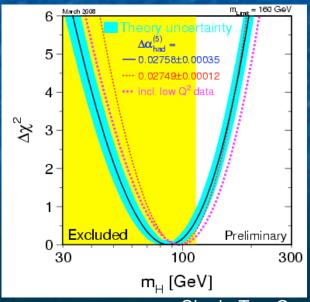


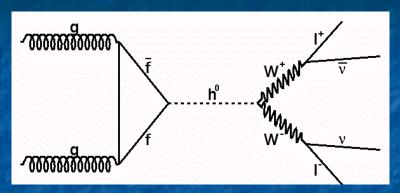
It's Not Looking Good For a Fourth Generation, However!

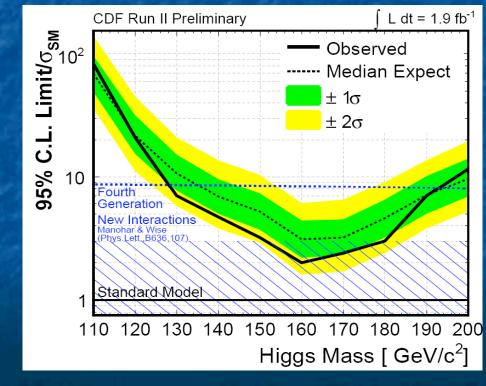
Higgs production via gluon-gluon fusion proceeds mostly via a top loop in a 3-gen model, and gets a boost from heavier quarks if they exist.



$$m_{\rm H}^{\rm EW-fit} = 87_{-27}^{+36} \,\, {\rm GeV}$$

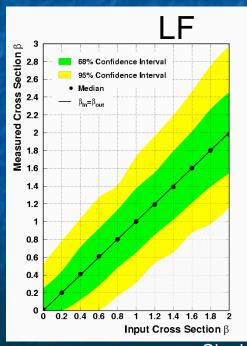




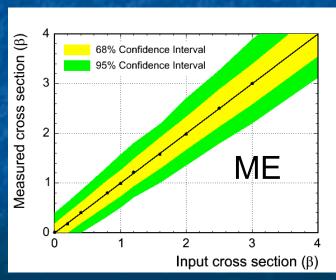


Cross Section Measurements

- Bayesian Technique selected
 - Flat prior in signal cross section $\sigma_{\text{s+t}}$
 - Integrate out rate and shape uncertainties
 - Check biases with pseudoexperiments with systematics fluctuated.
 - m_{top}=175 GeV assumed.



Linearity checked with systematically varied pseudoexperiments



Fermilab from the Air

Tevatron ring radius=1 km

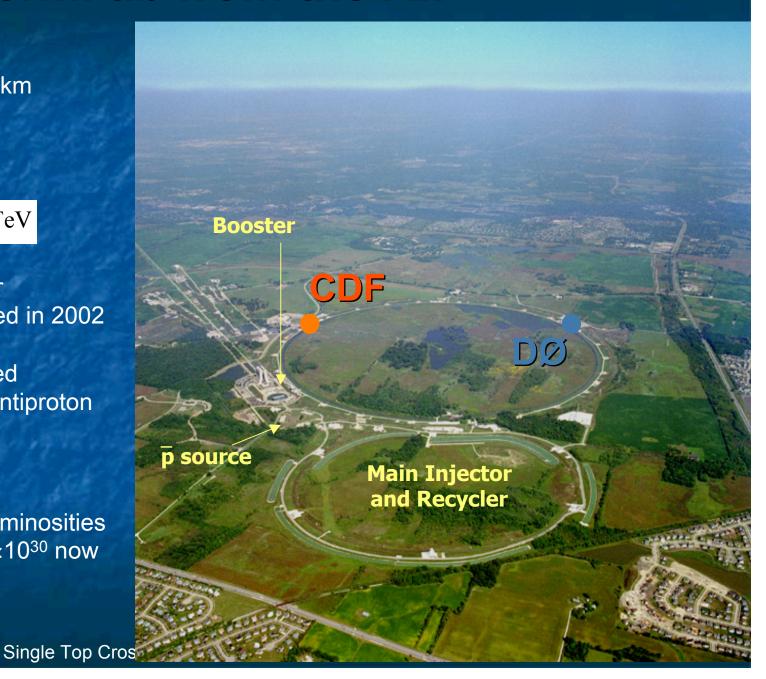
Protons on antiprotons

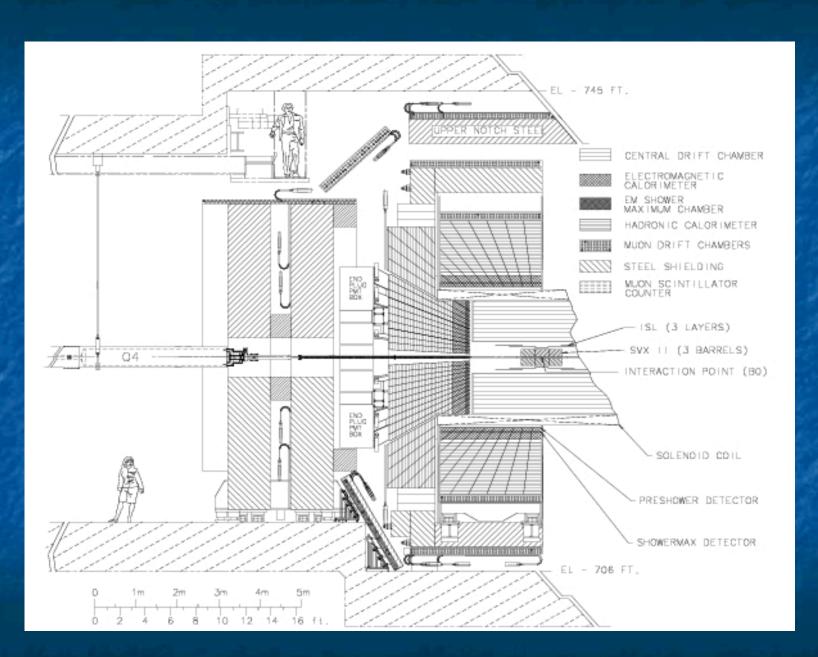
$$\sqrt{s_{p\bar{p}}} = 1.96 \text{ TeV}$$

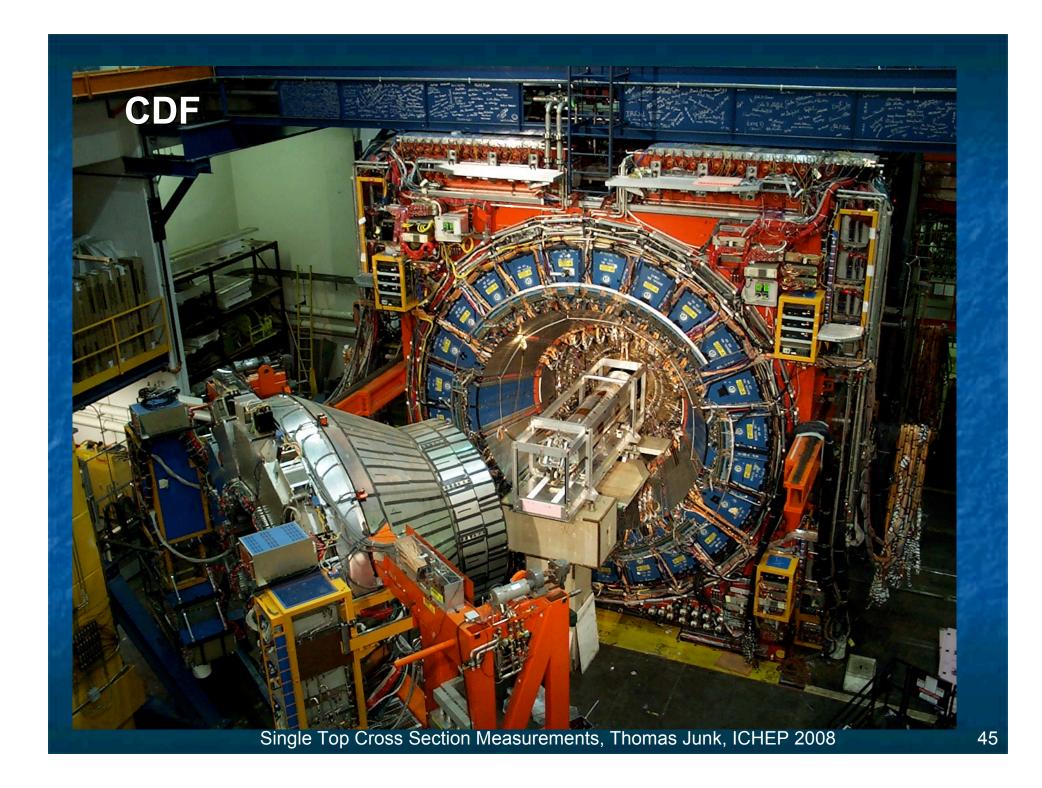
Main Injector commissioned in 2002

Recycler used as another antiproton accumulator

Start-of-store luminosities exceeding 250×10³⁰ now are routine

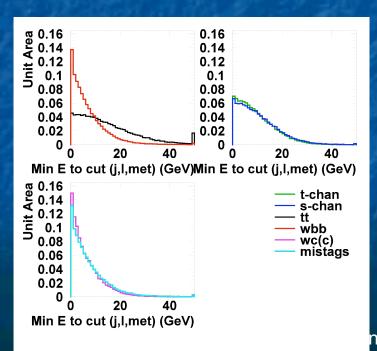


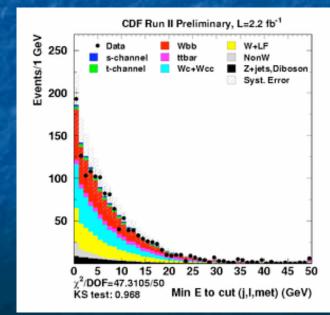




High-Significance Events Near Cuts

- Isolation removes QCD background, but signal likes to have the lepton close to a jet
- Jet energies and Missing Et have falling spectra, even for signal. Cuts are designed to keep within the validity region of the jet corrections
- t-channel signal throws one jet at very high η , challenging our ability to detect it. Geometrical detector acceptance and modeling of beam-splash and energy calibration set that cut.
- Lepton energy cut designed to get away from the 18-GeV trigger threshold





CDF BDT Variable list, 2-jet bin

			T -	
	2-jets, 1-tag		2-jets, 2 -tag	
Rank	Variable	Variable Importance	Variable	Variable Importance
1	KaNN	4.597e-01	mJ1J2	7.453e-02
2	mJ1J2	1.799e-01	Mlnub	6.599 e-02
3	QEta	1.077e-01	wmt	6.453 e - 02
4	Mlnub	4.684e-02	MetJ1DPhi	5.142e-02
5	$_{ m J1Et}$	4.570e-02	Mlnuj1j2	4.805e-02
6	wmt	1.890e-02	KaNN	4.656e-02
7	$_{ m J2Et}$	1.874e-02	cosLepJ1	4.645e-02
8	LepPt	1.639e-02	J1Eta	4.563e-02
9	cosLepJ1	1.572e-02	J1Et	4.522e-02
10	MetLepDPhi	1.481e-02	J2Et	4.424e-02
11	J2Eta	1.063e-02	LepEta	4.366e-02
12	LepEta	8.744e-03	MetJ2DPhi	4.339e-02
13	Mlnuj1j2	8.106e-03	LepJ2DPhi	4.252e-02
14	$_{ m Ht}$	7.755e-03	cosLepJ2	4.249 e-02
15	MetJ1DPhi	7.589e-03	MetLepDPhi	4.098e-02
16	LepJ1DPhi	6.232e-03	LepJ1DPhi	3.886e-02
17	\cos LepJ2	6.035e-03	LepPt	3.869 e-02
18	met	4.676e-03	QEta	3.850e-02
19	MetJ2DPhi	4.174e-03	met	3.840e-02
20	J1Eta	3.922e-03	J2Eta	3.573 e-02
21	WEta	3.881e-03	Ht	3.227e-02
22	LepJ2DPhi	3.839e-03	WEta	3.188e-02

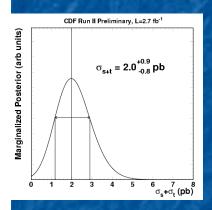
CDF BDT Variable List, 3-jet bin

	3-jets, 1-tag			3-jets, 2-tag	
Rank	Variable	Variable Importance	Variable	Variable Importance	
1	KaNN	3.846e-01	QEta	7.785e-02	
2	mJ2J3	7.124e-02	Ht	5.032e-02	
3	wmt	5.449e-02	Mlnub	4.698e-02	
4	mJ1J3	5.358e-02	mJ2J3	4.100e-02	
5	$_{ m Ht}$	4.978e-02	mJ1J2	3.864e-02	
6	QEta	4.729e-02	mJ1J2J3	3.735e-02	
7	Mlnub	3.924e-02	MetLepDPhi	3.677e-02	
8	mJ1J2J3	2.642e-02	wmt	3.632e-02	
9	$_{ m J3Et}$	2.609e-02	J1Eta	3.573e-02	
10	J2Eta	2.549e-02	cosLepJ1	3.447e-02	
11	mJ1J2	2.490e-02	J2Et	3.400e-02	
12	cosLepJ1	2.370e-02	J3Et	3.324e-02	
13	LepPt	1.910e-02	mJ1J3	3.307e-02	
14	LepEta	1.774e-02	J2Eta	3.275e-02	
15	MetLepDPhi	1.382e-02	cosLepJ3	3.046e-02	
16	J1Eta	1.373e-02	J1Et	3.017e-02	
17	LepJ1DPhi	1.223e-02	WEta	2.993e-02	
18	\cos LepJ2	1.069e-02	MetJ2DPhi	2.988e-02	
19	cosLepJ3	9.661e-03	LepJ1DPhi	2.898e-02	
20	J1Et	8.641e-03	KaNN	2.894e-02	
21	MetJ3DPhi	8.278e-03	LepEta	2.873e-02	
22	MetJ2DPhi	8.256e-03	LepJ3DPhi	2.855e-02	
23	met	8.141e-03	Mlnuj1j2j3	2.828e-02	
24	MetJ1DPhi	7.496e-03	MetJ3DPhi	2.692e-02	
25	LepJ3DPhi	6.743e-03	\cos LepJ2	2.616e-02	
26	J2Et	6.133e-03	LepPt	2.518e-02	
27	LepJ2DPhi	6.035e-03	MetJ1DPhi	2.508e-02	
28	WEta	5.932e-03	LepJ2DPhi	2.204e-02	
29	Mlnuj1j2	5.417e-03	met	2.160e-02	
30	Mlnuj1j2j3	5.108e-03	Mlnuj1j2	2.060e-02	

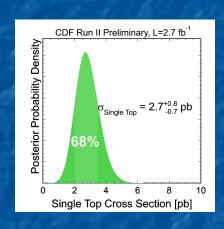
Cross Section Measurements

- Bayesian posterior marginalized over uncertain nuisance parameters
- Flat prior taken in $\sigma_s + \sigma_t$

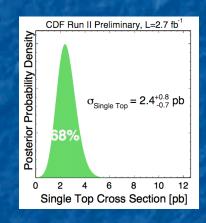




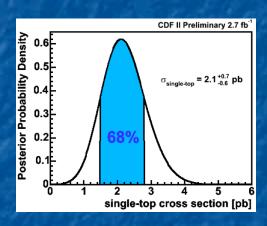
Likelihood Function



Matrix Element



Decision Trees



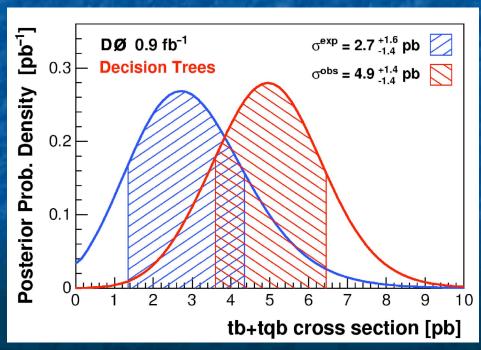
Neural Networks

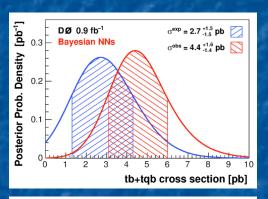
All Assuming m_t=175 GeV

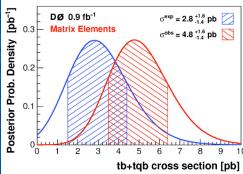


Cross Section Measurements

- Bayesian technique selected
- Flat prior in σ_s + σ_t
- Systematic Uncertainties integrated over (marginalized)







Three measurements are highly correlated between 60% and 70%

High-Score Events: CDF BDT > 0.6

